THE LANCET
Respiratory Medicine

Supplementary appendix

This appendix formed part of the original submission and has been peer reviewed. We post it as supplied by the authors.

Supplement to: Ramakrishnan R, Nicolau Jr DV, Langford B, et al. Inhaled budesonide in the treatment of early COVID-19 (STOIC): a phase 2, open-label, randomised controlled trial. *Lancet Respir Med* 2021; published online April 9. http://dx.doi.org/10.1016/S2213-2600(21)00160-0.

Supplementary Material

Inhaled budesonide in the treatment of early COVID-19 illness: a randomised controlled trial

Sanjay Ramakrishnan, MBBS[1,2,3+], Dan V. Nicolau Jr. PhD[1,4,5+], Beverly Langford RGN[1,2], Mahdi Mahdi BSc[1,2], Helen Jeffers RGN[1,2], Christine Mwasuku PGDip[1,2], Karolina Krassowska MSc[1,2], Robin Fox MBBS[6,7], Ian Binnian MBChB[8], Victoria Glover MBBS[9], Stephen Bright MSc[10], Christopher Butler FMedSci[11], Jennifer L Cane PhD[1,2], Andreas Halner BA[1], Philippa C Matthews DPhil[1,12], Louise E Donnelly PhD[13], Jodie L Simpson PhD[14], Jonathan R Baker PhD[13], Nabil T Fadai PhD[15], Stefan Peterson PhD[16], Thomas Bengtsson MSc [16] Peter J Barnes FRS[13], Richard EK Russell PhD[1,2,17], Mona Bafadhel PhD[1,2*]

+Joint first authors
*Senior and corresponding author

Affiliations:

1. Nuffield Department of Clinical Medicine, University of Oxford, United Kingdom
2. National Institute for Health Research (NIHR) Oxford Biomedical Research Centre (BRC), United Kingdom
3. School of Medical and Health Sciences, Edith Cowan University, Perth, Australia
4. UQ Centre for Clinical Research, The University of Queensland, Brisbane, Australia
5. School of Mathematical Sciences, Queensland University of Technology, Brisbane, Australia
6. Bicester Health Centre, Coker Close, Bicester, United Kingdom
7. National Institute for Health Research (NIHR) Thames Valley and South Midlands, United Kingdom
8. Eynsham Medical Group, Eynsham, United Kingdom
9. White Horse Medical Practice, Faringdon, United Kingdom
10. Windrush Medical Practice, Witney, United Kingdom
11. Nuffield Department of Primary Health Care Sciences, University of Oxford, United Kingdom
12. Department of Infectious Diseases and Microbiology, Oxford University Hospitals NHS Foundation Trust, John Radcliffe Hospital, Oxford, United Kingdom
13. National Heart and Lung Institute, Imperial College, London, United Kingdom
14. Priority Research Centre for Healthy Lungs, School of Medicine and Public Health, University of Newcastle, NSW Australia.
15. School of Mathematical Sciences, University of Nottingham, Nottingham, United Kingdom
16. STATMIND, Lund, Sweden
17. Southernhealth NHS Foundation Trust, Hampshire, UK.

Corresponding author: Prof Mona Bafadhel

Address: Respiratory Medicine Unit, NDM Research Building, Nuffield Department of Medicine, University of Oxford, Old Road Campus, Roosevelt Drive, Oxford, OX3 7FZ, United Kingdom

Tel: (+44) (0) 1865 612898 **Email:** mona.bafadhel@ndm.ox.ac.uk

Methods

SARS-CoV-2 detection from Nasopharyngeal swabs: Nasopharyngeal swabs were collected from all participants at randomisation, day 7 and day 14. Participants were asked to perform self-swabs, using one swab collecting sample from the nasopharynx and throat. Research nurses performed the nasopharyngeal swab if the participant was unable to. Swabs were stored upon collection immediately in RNAse solution, for SARS-CoV-2 deactivation and transferred in protected collection pots to the main Respiratory Medicine laboratory at the University of Oxford.

Virus extraction and quantification: RNA was extracted from clinical samples using the QIAamp Viral RNA Mini Kit (Catalogue: 52906, Qiagen) following manufacturer's instructions. 420 µl of sample was extracted and eluted into 40 µl Buffer AVE. 10 µl of eluted RNA was assayed using the Taqman fast virus 1-step master mix (ThermoFischer Scientific, Loughborough, UK), utilising oligonucleotide primers (600 nM forward and 800 nM reverse per reaction) and fluorescent conjugated probes (two probes 100 nM each) (see supplementary table 1 for sequences) (Eurofins Genomics, Wolverhampton, UK) for the detection of the viral RNase P gene (RdRP) gene region of SARS-CoV-19 using the ABI 7500 SDS Instruments (Applied Biosystems). Virologic testing for SARS-CoV-2 infection was performed by quantitative real-time RT-PCR (RT-qPCR). Assay data are presented in cycle threshold (units). The limit of detection was 40 cycles.

Sequences for primer and probes for RT-qPCR

Dual Labelled probe P2: CAG GTG GAA CCT CAT CAG GAG ATG C
Dual labelled probe P1: CCA GGT GGW ACR TCA TCM GGT GAT GC
PCR primer RDRP F: GTG ARA TGG TCA TGT GTG GCG G
PCR primer RDRP R: CAR ATG TTA AAS ACA CTA TTA GCA TA

SARS-CoV-2 antibody detection: All participant samples were collected in 5ml serum separator tubes manufactured by BD Vacutainer®. Samples were labelled using a pseudo anonymised code and sent to the Oxford University Hospital NHS Foundation Trust microbiology laboratory within 6 hours of sample collection. Serology for IgG to nucleocapsid protein was performed using the Abbott Architect i2000 chemiluminescent microparticle immunoassay (Abbott, Maidenhead, UK). Antibody levels ≥1.40 arbitrary units were considered positive.

NHS service set-up in UK: NHS service delivery in the UK during the pandemic included COVID virtual wards and home monitoring services, which included oxygen at home for COVID-19 infection. This meant that patients that were safe to discharge (including those that reached the primary outcome) were sent home with GP led monitoring and/or additional treatment (which could include analgesia, systemic corticosteroids, IV fluids and oxygen therapy) or specialist care monitoring in 'virtual' wards. Patients requiring significant respiratory support and/or not safe for discharge were admitted to hospital.

Statistical Analysis

See attached Statistical analysis plan for primary and secondary analyses.

Stochastic Simulations of a Virtual Twin of the STOIC Study: To further shed light on the data generated in the STOIC trial, we performed stochastic simulations of a "virtual" trial with the same design, primary endpoints and duration as STOIC. In our simulations, each of a number of virtual "patients" are recruited (R) assigned to either the BUD or UC arms, as illustrated in illustration 1. The simulation visits each virtual patient once per simulation day. The patient may stay in the BUD or UC "state" or transition to drop-out (DO), recovery (RES) or reaching primary end-point (PO), with probabilities for each transition pre-set, as illustrated (and summing to 1). A random number between 0 and 1 will be generated on each such day and the decision which transition to make is made in accordance to its value.

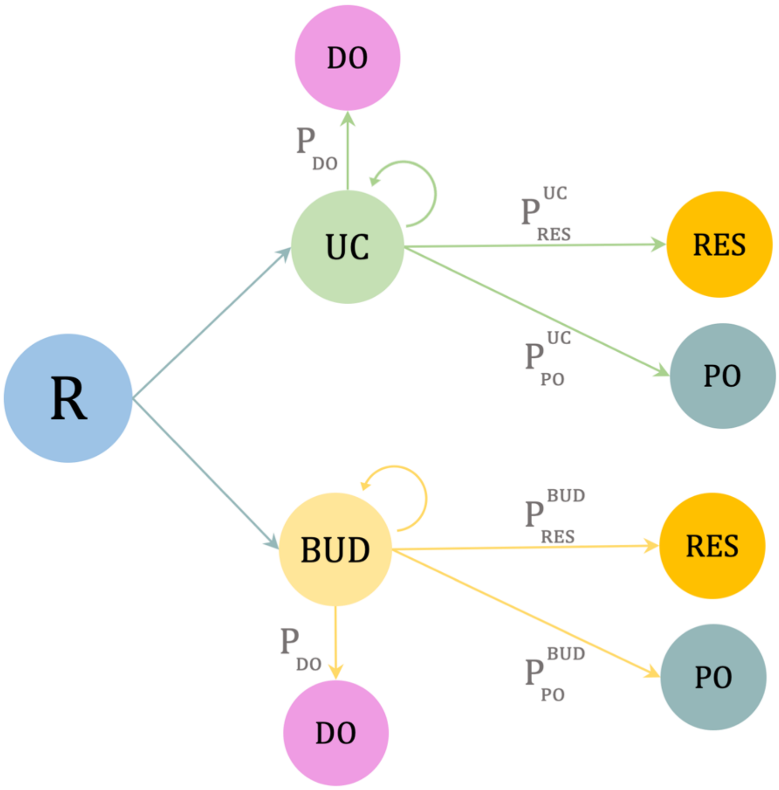


Illustration 1

Patients are recruited (R) to either the BUD or UC arms and may, during each day of the virtual trial either remain recruited or have symptom resolution (RES), reach primary outcome (PO) or withdraw from the study (DO), according to probabilities, as shown.

In particular, the ratio of P^{UC}_{PO} to P^{BUD}_{PO} represents a treatment effect, the purported reduction in the odds (on a daily basis) of reaching primary outcome that can be attributed to the effect of the budesonide treatment. We parametrised our model for the same number of patients recruited to both arms as in STOIC and worked backwards to estimate the maximum-likelihood daily probabilities for each of the transitions such that the mean virtual outcomes for the UC arm are the same, on average, as our findings (computer code available on request). We then studied the relationship between the ratio of P^{UC}_{PO} to P^{BUD}_{PO} and that of the ratios of patients reaching primary outcome during the virtual trial for the two arms. The results are illustrated in the main paper, Figure 5, with the grey envelope around the yellow curve (mean), representing the 95% confidence interval of outcomes. They indicate that in order to reach an average ratio of 10:1, a daily reduction in the odds of reaching primary outcome is approximately 3000%, with a minimum value (at 95% confidence) of approximately 400%, confirming that a very large daily treatment effect can be attributed to the use of budesonide inhalers for COVID-19 patients.

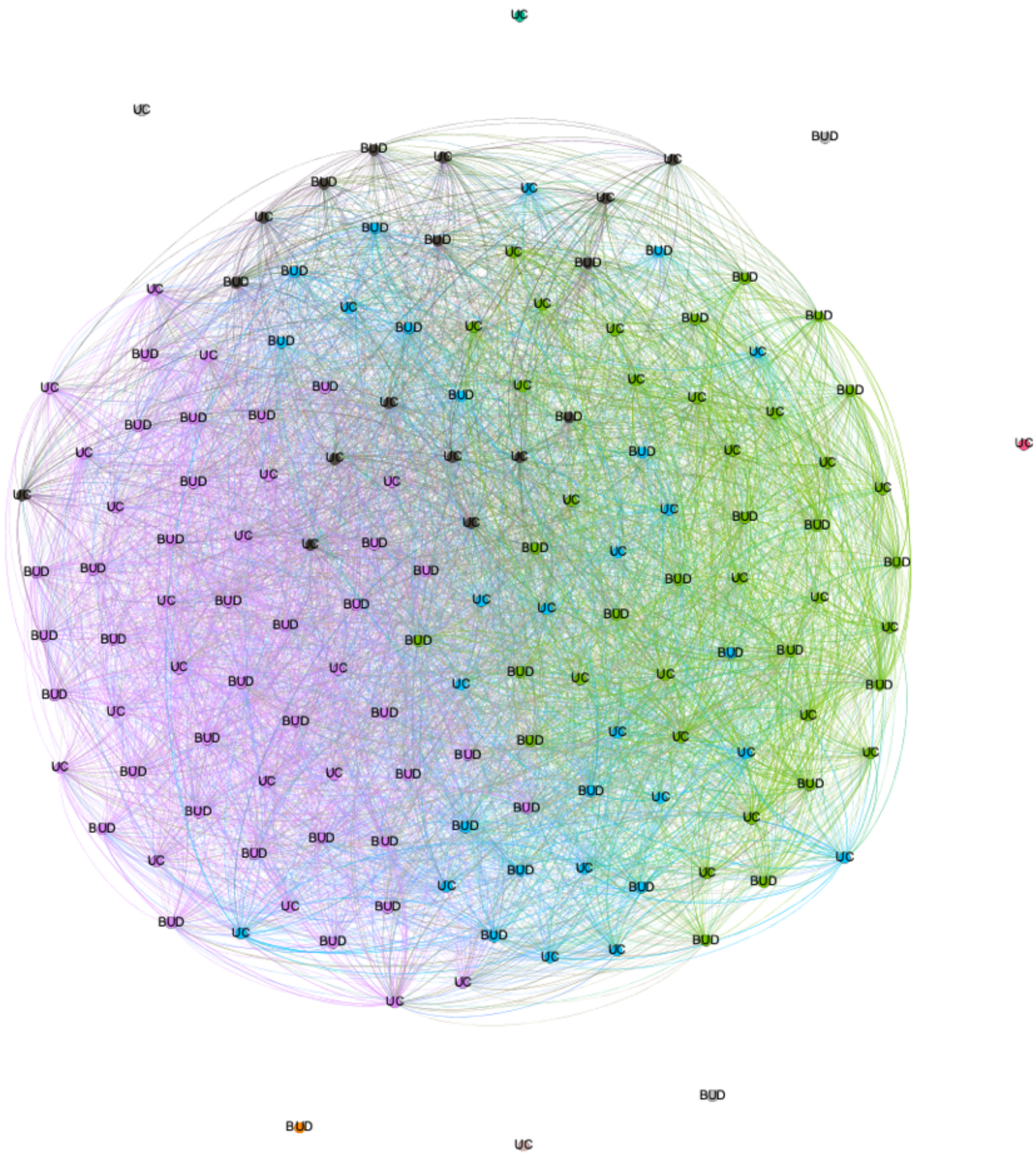
***Community detection algorithms*:** To further elucidate the clinical trajectories of patients in the two arms of the STOIC study, we used community detection methods to interpret our finding as complex networks. Specifically, we treat each patient as a node in a network. The edges between nodes are averages of correlations between their time series in respect of a) highest daily temperature, b) lowest daily temperature, c) highest oxygen saturations, d) lowest oxygen saturations and e) heart rate. We weighted each of these time series correlations equally and only used data from days 1-14 because after this point, the data entries were sparser and yielded poor-quality correlation values.

In order to detect "communities" of patients that behave similarly to each other and differently to others, if these exist, we computed, over the resulting network, whose edges between two nodes *i* and *j* are denoted by A_{ij}, the maximum of a function of the type

$$Q(\vec{\sigma}) = \frac{1}{A_{\text{tot}}} \sum_{i,j} [A_{ij} - \langle A_{ij} \rangle] \delta(\sigma_i, \sigma_j)$$

Where Q is the network modularity and δ indicates module membership, i.e. it is equal to 1 is two nodes belong to the same module. A_{tot} is the sum over all the edge strengths computed as described above. This approach reveals a complex structure within our patient network, comprised of 4 modules, corresponding broadly to recovery trajectories (poor or effective) and each dominated by patients from either the BUD or UC arms (illustration 2).

This analysis is unique in that it looks at patient parameters over time in context, rather than at each in isolation. The results confirm qualitatively that the patients in the BUD arm recovered better in the totality of their measured parameters than did those in the UC arm, despite the small scale of our study. Isolated nodes indicate patients with sparsely recorded data or that dropped out and so did not fit in any of the communities



Supplementary Table S1. Demographics and clinical characteristics of study participants in the intention to treat population at study enrolment

Characteristic	**Budesonide (n = 73)**	**Usual care (n = 73)€**
Age, years#	44 (19-71)	45 (19-79)
Female sex, no. (%)	41 (56%)	43 (61%)
White ethnicity, no. (%)	67 (92%)	66 (93%)
Body mass index, kg/m^2	26 (5.1)	26 (4.5)
Number of co-morbidities, no. ¥	1 (0-2)	1 (0-1)
Duration of symptoms prior to randomisation, days¥	3 (2-5)	3 (2-4)
Evidence of COVID positive status, no. (%)	69 (96)	68 (93)
Presence of symptoms at baseline, no. (%)		
Cough	58 (80%)	49 (70%)
Fever	52 (71%)	45 (63%)
Headache	41 (56%)	38 (54%)
Fatigue	33 (45%)	24 (34%)
Loss of sense of smell/taste	25 (34%)	30 (42%)
Gastrointestinal symptoms	12 (16%)	12 (17%)
Breathlessness	11 (15%)	11 (16%)
Myalgia	7 (10%)	10 (14%)
Nasal symptoms	3 (4%)	5 (7%)
Sore throat	0 (0%)	2 (3%)
Chest pain/tightness	4 (5%)	1 (1%)
Other	7 (10%)	8 (11%)
FLUPRO score*	0.82 (0.49)	0.81 (0.44)
CCQ score*	0.74 (0.44)	0.67 (0.41)
Highest temperature recorded*¥ in degrees centigrade	36.6 (35.2-39.0)	36.6 (35.5-38.3)
Lowest Oxygenation recorded*¥ as % saturation	96 (95-97)	96 (95-97)

Data presented as mean (SD) unless stated; #mean (range); *at randomisation; ¥Median (IQR); FLUPRO InFLUenza Patient Reported Outcome questionnaire; CCQ Common Cold Questionnaire; € 2 participants withdrew from study after study randomisation and only gender, age, and COVID-19 infection status were collected.

Supplementary Table S2: Demographics and clinical characteristics of study participants with a primary outcome compared to participants with symptom resolution in the per-protocol population

Characteristic	**Primary outcome (n = 11)**	**Achieved symptom resolution (n = 128)**	**P value**
Age, years	45 (19-79)	45 (24-57)	0.90
Female sex, no. (%)	9 (82%)	71 (55%)	0.09
White ethnicity, no. (%)	11 (100%)	118 (92%)	0.31
Body mass index, kg/m^2	27 (4.5)	26 (7.0)	0.70
Number of co-morbidities, no. ¥	1 (0-1)	1 (0-2)	0.90
Duration of symptoms prior to randomisation, days¥	3 (2-4)	3 (2-4)	0.56
Evidence of COVID positive status, no. (%)	9 (90)	120 (93)	0.54
Presence of symptoms at baseline, no. (%)			
Cough	8 (73%)	96 (74%)	0.90
Fever	8 (73%)	86 (67%)	0.68
Headache	8 (73%)	70 (54%)	0.24
Fatigue	3 (27%)	52 (40%)	0.40
Loss of sense of smell/taste	4 (36%)	51 (40%)	0.84
Gastrointestinal symptoms	2 (18%)	22 (17%)	0.92
Breathlessness	3 (27%)	19 (15%)	0.27
Myalgia	2 (18%)	14 (11%)	0.46
Nasal symptoms	1 (9%)	7 (5%)	0.62
Sore throat	1 (9%)	1 (1%)	0.026
Chest pain/tightness	0 (0%)	5 (4%)	0.51
Other	2 (18%)	13 (10%)	0.40
FLUPRO score*	0.93 (0.33)	0.81 (0.47)	0.28
CCQ score*	0.80 (0.31)	0.70 (0.44)	0.35
Highest temperature recorded*¥ in degrees centigrade	36.6 (35.2–39.0)	36.7 (35.8-37.2)	0.87
Lowest Oxygenation recorded*¥ as % saturation	96 (84-99)	96 (93-99)	0.80

Data presented as mean (SD) unless stated; #mean (range); *at randomisation ¥Median (IQR); FLUPRO InFLUenza Patient Reported Outcome questionnaire; CCQ Common Cold Questionnaire

Supplementary Table S3. Delta mean change in FLUPro® symptoms between days 0 and 14 for the individual domains in the BUD and UC study arms.

FLUPRO domain	Budesonide	Usual care	P value*
Systemic	-0.94	-0.80	0.034
Nose	-0.72	-0.56	0.093
Chest/Respiratory	-0.48	-0.37	0.165
Eyes	-0.28	-0.23	0.325
Throat	-0.61	-0.57	0.542
Gastrointestinal	-0.30	-0.30	0.973

*from the ANCOVA model adjusting for treatment, age (>40, ≤40), sex, no. of comorbidities (≥2, ≤1) and baseline

Supplement Figure Legends

Figure S1. Sensitivity analysis for time to clinical recovery in patients with confirmed SARS-CoV-2 infection

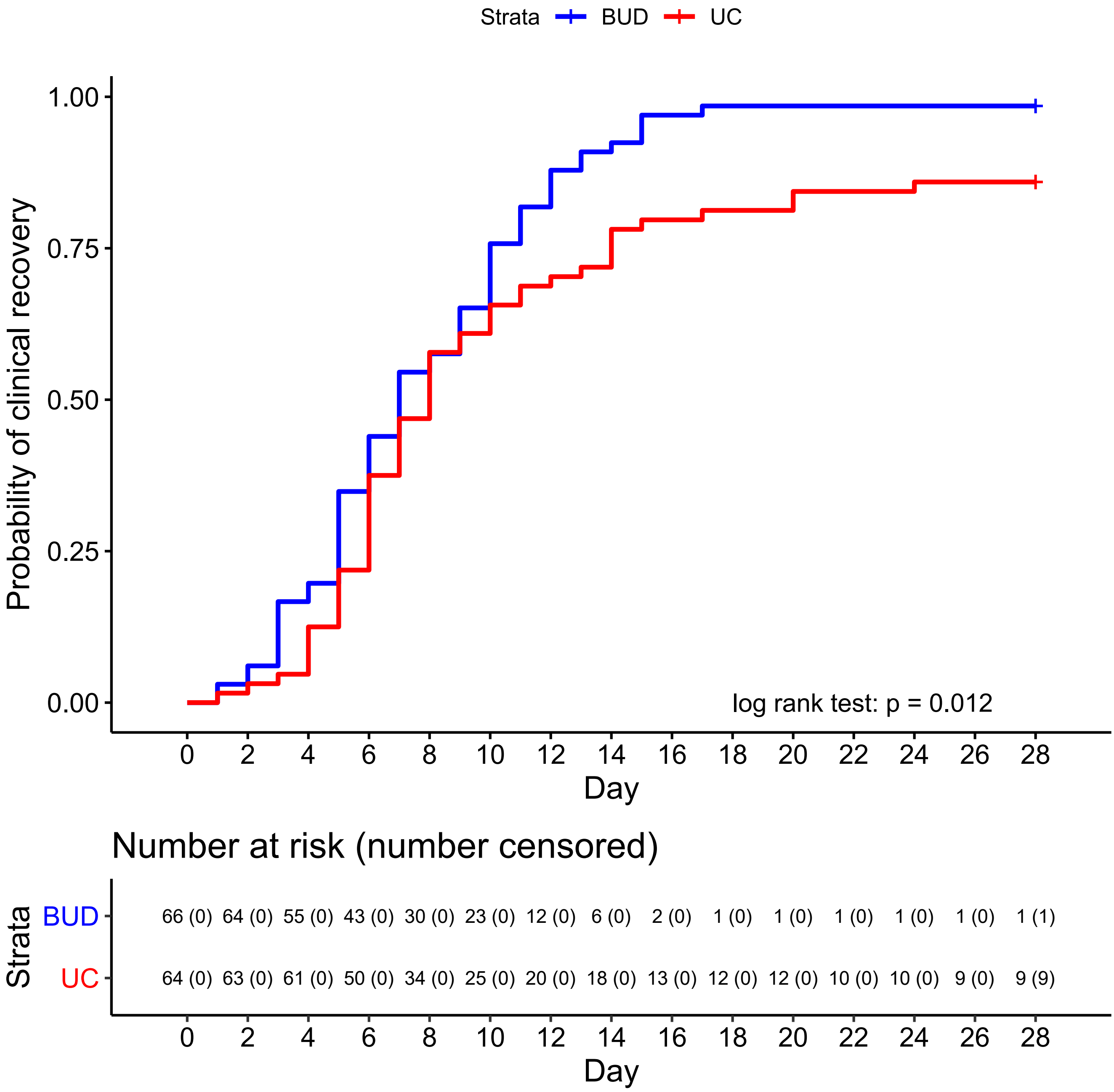
Figure S2. Daily peak temperature in BUD and UC participants. Trends indicate that daily highest temperature fell more rapidly in the BUD (-0.113 degrees Celsius per day) than the UC (-0.096 degrees Celsius per day) arm.

Figure S3. Time to symptom resolution as measured by the FLUPro®

Figure S4. Daily mean CCQ score for BUD and UC arm over 14 days. Vertical bars indicate standard error.

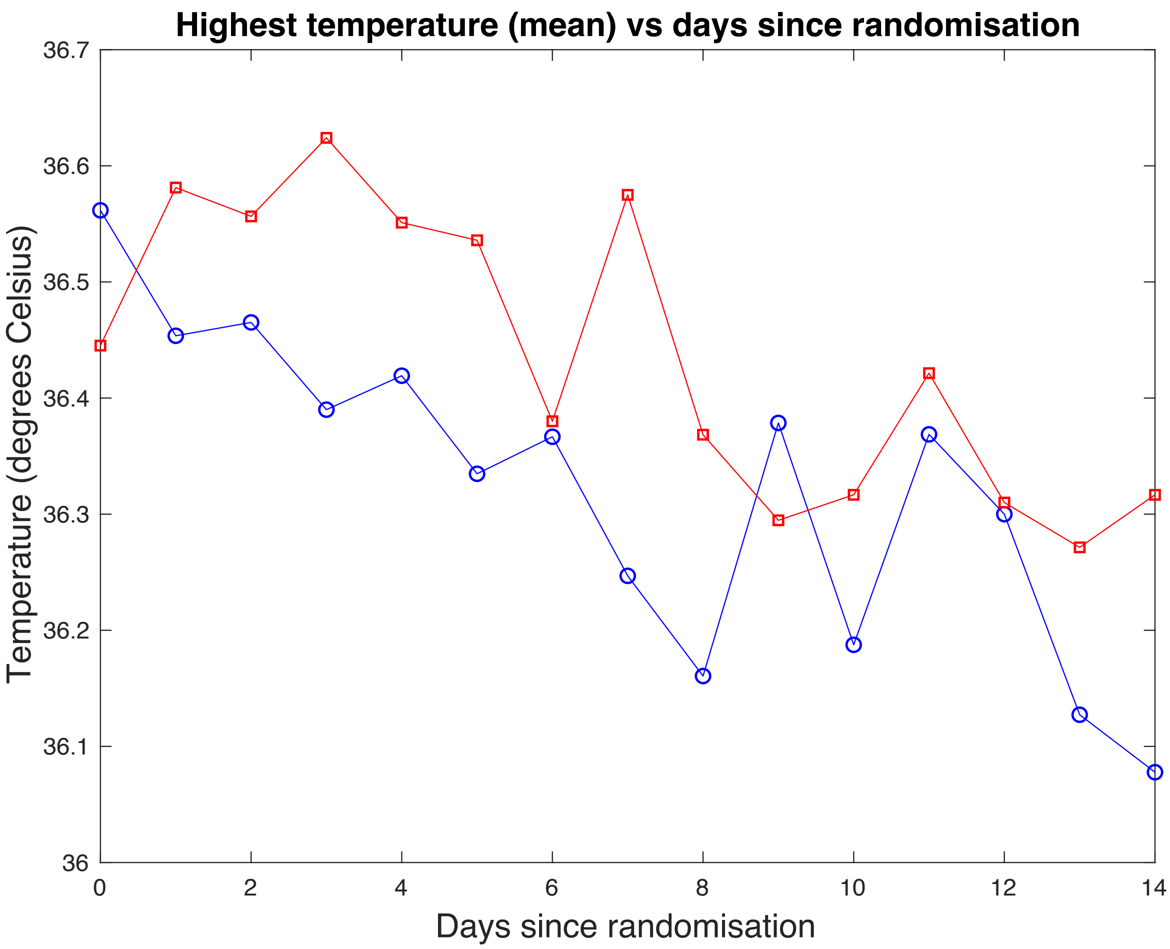
Figure S5. Violin plots presenting cycle threshold (CT) over 3 study visits (day 0, 7 and 14) in the BUD and UC arm. Solid line represents median, dashed lines represent upper and lower interquartile. BUD = budesonide; UC = usual care. Lower limit of detection set at CT 40, CT values above 40 indicate undetectable virus

Figure S1



Strata
BUD
UC
Probability of clinical recovery
1.00
0.75
0.50
0.25
0.00
log rank test: p = 0.012
0 2 4 6 8 10 12 14 16 18 20 22 24 26 28
Day
Number at risk (number censored)
Strata
BUD
UC
66 (0) 64 (0) 55 (0) 43 (0) 30 (0) 23 (0) 12 (0) 6 (0) 2 (0) 1 (0) 1 (0) 1 (0) 1 (0) 1 (0) 1 (1)
64 (0) 63 (0) 61 (0) 50 (0) 34 (0) 25 (0) 20 (0) 18 (0) 13 (0) 12 (0) 12 (0) 10 (0) 10 (0) 9 (0) 9 (9)
0 2 4 6 8 10 12 14 16 18 20 22 24 26 28
Day

Figure S2



Highest temperature (mean) vs days since randomisation
36.7
36.6
36.5
36.4
36.3
36.2
36.1
36
Temperature (degrees Celsius)
0
2
4
6
8
10
12
14
Days since randomisation

Figures S3

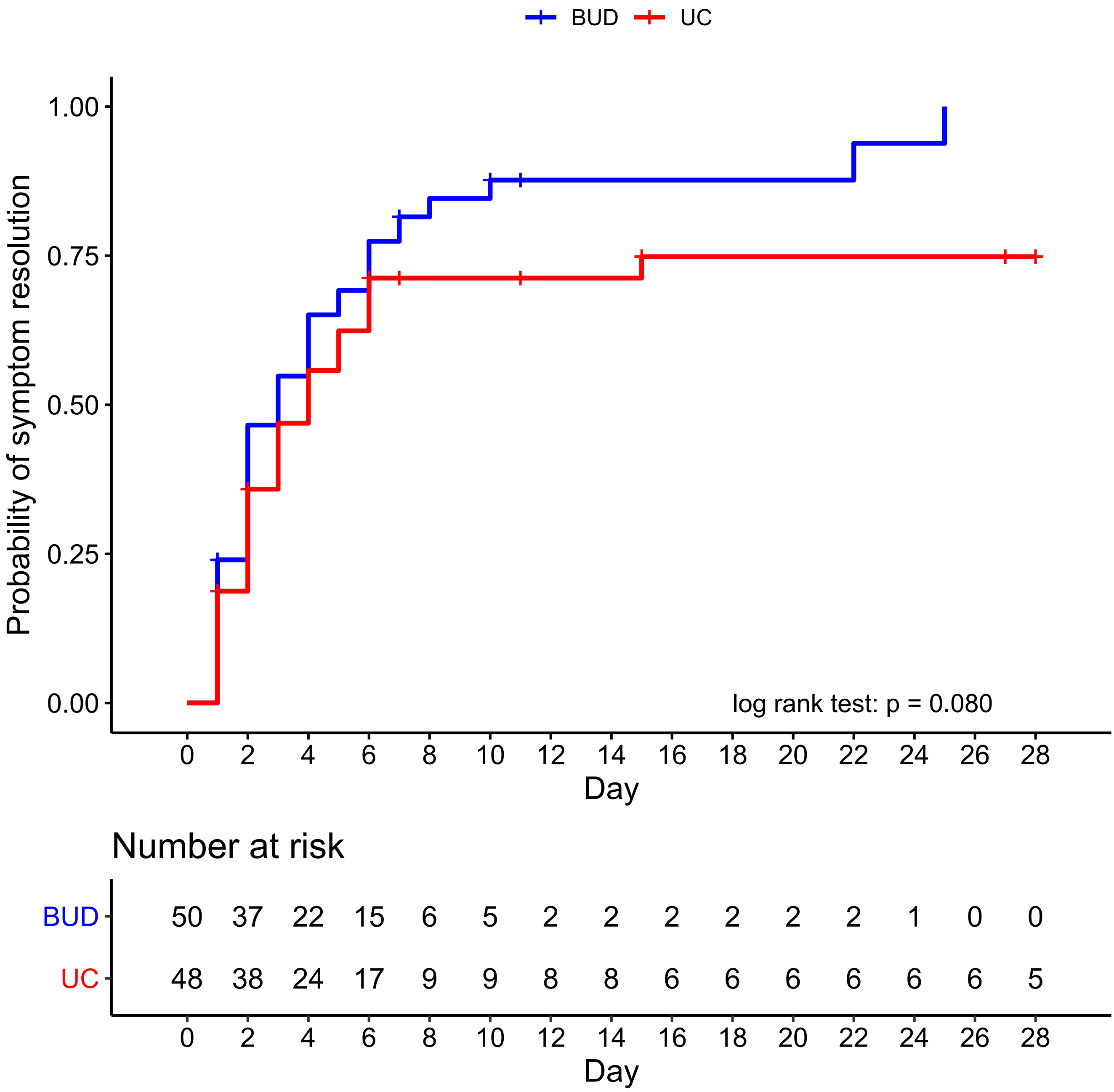
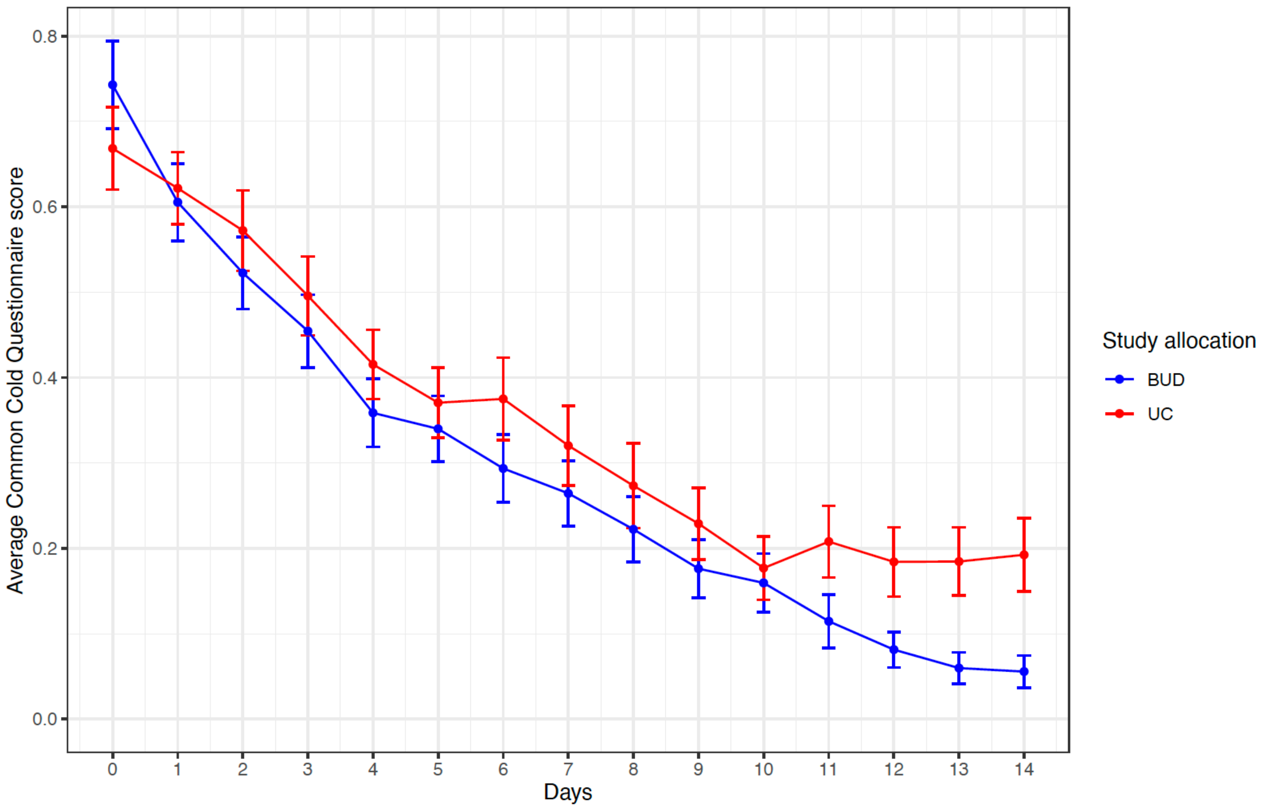
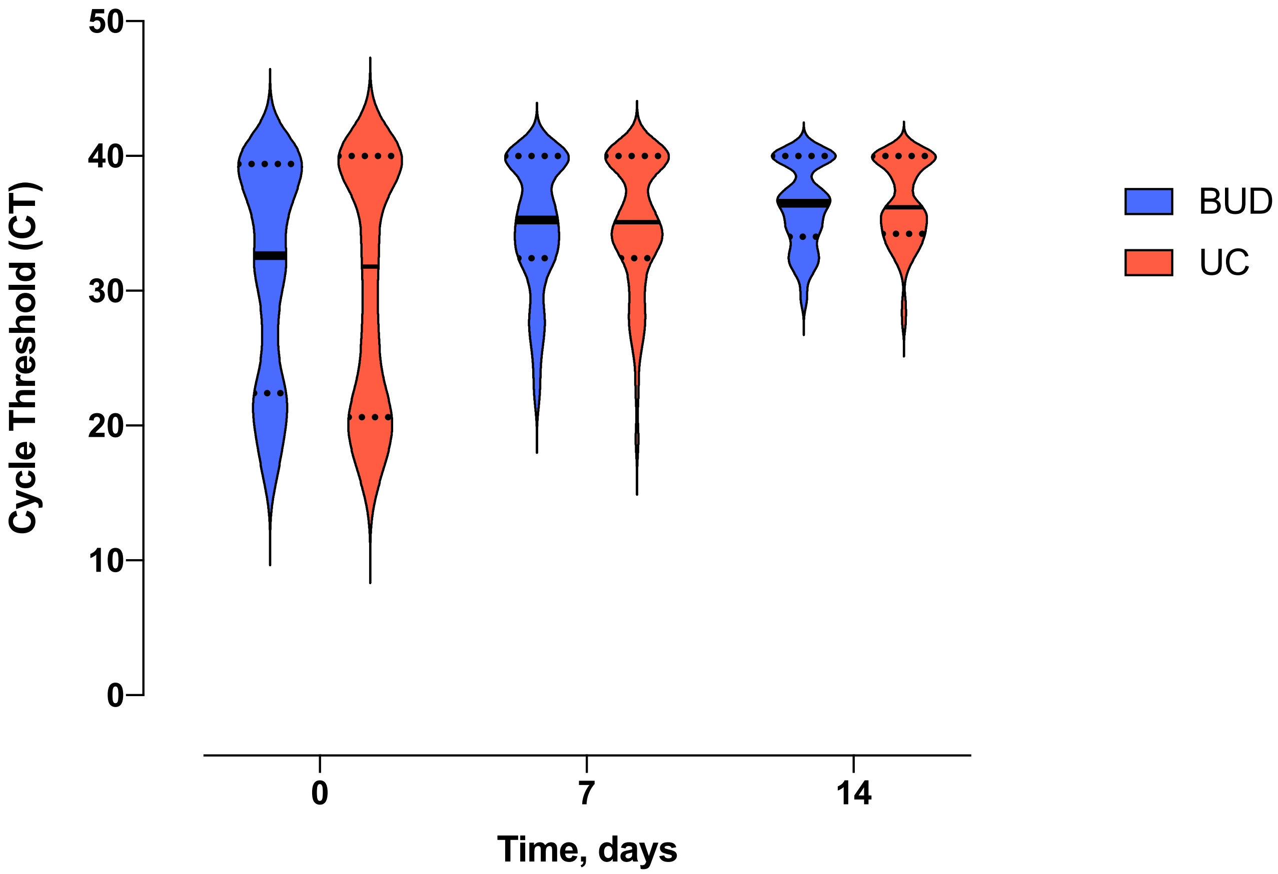

BUD
UC
Probability of symptom resolution
1.00
0.75
0.50
0.25
0.00
log rank test: p = 0.080
0 2 4 6 8 10 12 14 16 18 20 22 24 26 28
Day
Number at risk
BUD 50 37 22 15 6 5 2 2 2 2 2 2 1 0 0
UC 48 38 24 17 9 9 8 8 6 6 6 6 6 6 5
0 2 4 6 8 10 12 14 16 18 20 22 24 26 28
Day

Figure S4



day	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Bud	67	66	67	66	67	66	68	67	68	68	68	68	66	67	68
UC	67	66	66	63	66	66	65	66	67	66	68	68	68	66	68

Figure S5



Treating COVID-19 infections with inhaled corticosteroids (The STOIC Trial)

Statistical Analysis Plan

Version 2.0 – 09/Dec/2020

Based on Protocol version 2.0 – 26May2020

Trial registration: NCT04416399

Contents

Introduction

This document details the proposed data presentation and analysis for the main paper(s) and final study reports from the **Treating COVID-19 infections with inhaled corticosteroids (The STOIC Trial).**

The results reported in these papers should follow the strategy set out here. Subsequent analyses of a more exploratory nature will not be bound by this strategy, though they are expected to follow the broad principles laid down here. The principles are not intended to curtail exploratory analysis (for example, to decide cut-points for categorisation of continuous variables), nor to prohibit accepted practices (for example, data transformation prior to analysis), but they are intended to establish the rules that will be followed, as closely as possible, when analysing and reporting the trial.

The analysis strategy will be available on request when the principal papers are submitted for publication in a journal. Suggestions for subsequent analyses by journal editors or referees, will be considered carefully, and carried out as far as possible in line with the principles of this analysis strategy; if reported, the source of the suggestion will be acknowledged.

Any deviations from the statistical analysis plan will be described and justified in the final report of the trial. The analysis should be carried out by an identified, appropriately qualified and experienced statistician, who should ensure the integrity of the data during their processing. Examples of such procedures include quality control and evaluation procedures.

Changes from previous version of SAP

A summary of key changes from earlier versions of SAP, with particular relevance to protocol changes that have an impact on the design, definition, sample size, data quality/collection and analysis of the outcomes will be provided. Include protocol version number and date.

Version number **Issue date**	**Author of this issue**	**Protocol Version & Issue date**	**Significant changes from previous version together with reasons**
2.0 09 Dec 2020	S Ramakrishnan		Stop analysis a priori detail described

BACKGROUND and Objectives

Background and rationale

During this CoVID-19 pandemic, there are no effective treatments for mild disease. To allow for better management of health resources during this pandemic, effective interventions to reduce the number of emergency department visits and hospitalisations are important The global evidence that patients with respiratory co-morbidity are not commonly found in patients with severe COVID-19 infection, implies that inhaled corticosteroids (ICS) have an effect. ICS are accessible, cheap, safe, and widely prescribed. If effective, this would be the ideal class of medication to prescribe to reduce COVID-19 related morbidity and mortality.

During the coronavirus global pandemic, it is felt that respiratory symptoms with a new cough, and/or fever and/or flu symptoms are going to be related to SARS-CoV-2 infection and thus consistent with COVID-19. Furthermore, the sensitivity of the PCR testing is approximately 70% and significantly affected by testing practices. In this study design, a significant proportion will be self-swabbed, such that we believe that the sensitivity will reduce to 50% at best. This means there would be a risk of providing false negative results to potential participants and therefore providing false reassurance that they are not required to self-isolate due to their symptoms (which is a requirement as per Public Health Guidance). Therefore, a positive COVID-19 test result is not a requirement of entry into the study; but the use of symptoms suggestive of, viral titres, PCR test, or antibodies will improve the external validity of the trial data.

The STOIC study will be an open label, randomised controlled trial comparing, an inhaled corticosteroid against standard of care. Participants randomised to the study will be recruited within 7 days of symptom onset. Participants in the inhaled corticosteroid arm will be given 1600 mcg of inhaled budesonide daily. Thus, this form of study design allows for recruitment of patients early in the course of the disease and give enough time for the intervention to be effective

Objectives

Objectives	Outcome Measures	Timepoint(s) of evaluation of this outcome measure (if applicable)
Primary Evaluate the efficacy of ICS therapy compared to standard care in participants with early CoVID-19 illness	1.Hospitalisation or emergency department attendance related to COVID	Day 0 to 28
Secondary Evaluate the effect of ICS therapy on physiology, symptoms, and viral load, compared to standard care in participants with early CoVID-19 illness	1. Body temperature 2. Blood oxygen saturation level	Day 0 to 28 Day 0 to 28

	3. Nasal/throat swab SARS-CoV-2 viral load 4. Common cold questionnaire and FluPRO questionnaire	Day 0, 7 and 14 Day 0 to 28
Exploratory Evaluate the effect of ICS therapy on whole blood, nasal and viral mediator responses compared to standard care in participants with early COVID-19 illness.	1. Nasal and viral mediator responses 2. Whole blood mediator responses	Day 0 and 14 Day 0 and 28

Study METHODS

Trial Design/framework

The STOIC trial is a randomised, open label, parallel group phase 2 superiority clinical trial conducted in the community in Oxfordshire, United Kingdom.

Randomisation and Blinding

Randomisation by minimisation for gender, age (18-40 or >40) and presence of co-morbidities (1 or less or >1) will be performed. Randomisation sequence will be created prior to study commencement using Microsoft Excel and the random number generation function. The randomisation sequence will be held in a secure network location at the University of Oxford. This study is unblinded and the first randomised patient was recruited on the 16th of July 2020.

Sample Size

At time of study inception, we assumed that 20% of all COVID-19 illness is severe and will require hospitalisation needing respiratory support. Using 80% power at 0.05 level, 199 patients in each arm are required to demonstrate a 50% reduction of hospitalisations (from 20% to 10%).

Statistical Interim Analysis, Data Review and Stopping guidelines

No interim analyses were planned at study commencement. Due to the well understood safety profile of the intervention, inhaled budesonide, a Data and Safety Monitoring Committee (DSMC) was not planned or convened.

On the 09th of December 2020, the Chief Investigator, requested an external and independent statistical consulting agency (StatMind®). This was as a result of the Chief Investigator being concerned about participant recruitment making study completion by February 2020 unlikely. The following concerns were raised

i. Two attempts at obtaining National Priority Support to extend to outside of one region were declined in March 2020 and September 2020 and thus STOIC recruitment was limited to one region.
ii. Reduced number of staff resources available to recruit, due to relocation.

iii. The waxing and waning nature of recruitment as a consequence of national lockdown strategies affecting the region.
iv. The successful emergence of vaccines targeting SARS-CoV-2 and implemented nationally
v. Adjunct competing studies assessing the same intervention with randomisation to inhaled budesonide starting on the 3rd of December 2020.

The complete communication with the statistician to determine if there is sufficient rigour to stop the study guideline is listed in Appendix A. Criteria to fulfil this was created *a priori* to be assessed by independent statistical consultants to determine if the study should continue recruitment or stop recruitment.

Timing of Final Analysis

All primary and secondary outcomes will be analysed at study stop. Exploratory outcomes will be analysed after the biological samples are analysed.

Blinded analysis

No blinded analysis will be undertaken.

Statistical Analysis Outline

Descriptive statistics will be used to describe variables between the groups in the interventional arm and the standard care arm. Appropriate parametric or non-parametric statistical tests will be performed. For continuous variables, the difference between treatments in the means or medians and the corresponding 95% confidence interval will be reported. For continuous variables, ANCOVA models (t-tests) with adjustment for the stratification of the study or Wilcoxon rank sum tests will be applied to compare the intervention and observational group. Missing data will be handled by last value caried forward (LOCF). For categorical variables, including the primary outcome, the number (and percentage) of patients in each category will be reported for each treatment group and chi-squared tests will be used for comparing treatment groups. Time to clinical recovery was illustrated using the Kaplan Meier method and median time given. Comparison was performed with a log rank test where participants with primary outcome was censored at Day 28. Characteristics of response will be identified following generation of receiver operator characteristic curves and correlation coefficients. As necessary sensitivity analysis will be performed.

Statistical Principles

Statistical Significance

All tests will be completed at a 5% 2-sided significance level. All comparative outcomes will be presented as summary statistics with 95% confidence intervals and reported in accordance with the CONSORT Statement (http://www.consort-statement.org). P-values will be reported to a minimum of 3 decimal places or as required by specific journals for publication.

Definition of Analysis Populations

Intent to treat (ITT): all participants included in their randomised groups.

Per protocol (PP): all participants who completed all scheduled study visits (either in person or telephone).

Corona positive (CP): all participants in PP with a conformed COVID-19 infection.

trial population and descriptive analyses

Representativeness of Study Sample and Patient Throughput

The flow of participants through each stage of the trial, including numbers of participants randomly assigned, receiving intended treatment, completing the study protocol, and analysed for the primary outcome will be provided in accordance with the CONSORT statement. Protocol violations/deviations and information relating to the number of participants screened, refused to participate (with reasons where available) will also be included.

Withdrawal from treatment and/or follow-up

Withdrawals/loss to follow-up together with reasons will be reported by intervention arm. To ensure that there are no differential losses between the groups this will be tested using absolute risk differences (95% confidence interval [CI]) and a chi-squared test. Any deaths (and their causes) will be reported separately.

Baseline Comparability of Randomised Groups

Baseline characteristics are reported by treatment group, including the stratification factors (as appropriate) with prognostic, demographic, and clinical covariates. Numbers (with percentages) for binary and categorical variables and means (and standard deviations), or medians (with lower and upper quartiles) for continuous variables will be presented; there will be no tests of statistical significance nor confidence intervals for differences between randomised groups on any baseline variable.

Description of Compliance with Intervention

Investigational medicinal product adherence will be recorded and reported in the participant characteristic table.

Reliability

The primary outcome is a hard measure of health care resource utilisation. The secondary outcomes used are clinically valid physiological measurements or validated symptom questionnaires.

Analysis

Outcome Definitions

Primary Outcome

Primary outcome measure: Urgent care, emergency department visit or hospitalisation for COVID-19. This definition includes all unplanned care for a COVID-19 related symptom/health problem for the 28-day period of the study. This outcome will be assessed daily and at each research visit.

Secondary outcomes

Clinical recovery

- Time to participant self-report symptom resolution. This is collected at the study visits, asked by the research nurse or during the daily calls. If, this date is absent, then this will be replaced by the date as to which the daily phone calls ceased. For participants that had a primary outcome this will be censored at 28 days.

Symptoms

- InFLUenza Patient Reported Outcome (FLUPRO®) – This validated symptom tool[1,2] will be used to capture participant symptoms until symptom resolution. Scores range from 0-4; where scores below 1 indicate symptom resolution. Handling of missing data and classification of domains will be calculated as per the FLUPRO® manual.

- Common Cold Questionnaire (CCQ) – This validated symptom tool[3] will be used to capture participant symptoms until symptom resolution. A higher score signifies worse symptoms. Scores range from 0-3, where scores of 0 signify no symptoms.

Physiology

- Blood oxygen saturations is defined as the fraction of oxygen saturated haemoglobin relative to total haemoglobin. Blood oxygen saturation is measured daily at home using a calibrated pulse oximeter. The highest and lowest blood oxygen saturation was recorded. Blood oxygen saturations less than 95% is categorised as abnormal.

- Body temperature is measured using a commercially available thermometer. The highest and lowest body temperature was measured daily by participants. Body temperature greater than 37.5°C was defined as a fever.

SARS-CoV-2 viral load on nasopharyngeal swab

- SARS-CoV-2 viral load will be measured using polymerase chain reaction (PCR). This will be measured on day 0, 7 and 14. The lower limit of normal (sensitivitiy of the assay) will be pre-definded as ½ the lowest detected standard curve.

Analysis Methods

Primary Outcome

Proportions of participants in each study arm who required urgent care, emergency department visit, or hospitalisation will be compared using a Chi squared test in a per-protocol analysis. Sensitivity analysis for COVID-19 infection will also be performed. An ITT analysis will also be performed.

Secondary outcomes

Symptoms – Time to clinical recovery

Time to clinical recovery will be assessed between the intervention and usual care group using Kaplan-Meier survival curves and a log-rank test. A univariate Cox proportional hazards analysis will be used to estimate the hazard ratio of the treatment effect together with a 95% confidence interval. For participants that have a primary outcome, censoring at 28 days will be performed (primary analysis).

Symptoms – FluPRO questionnaire

Th FLUPRO questionnaire is 32 item questionnaire which assess systems across 6 domains, with a score ranging from 0-4 for each question. The domains are nose, throat, eyes, chest/respiratory, gastrointestinal and body/systemic (see appendix B for details). The arithmetic mean of all items will be calculated. A score <=1.0 is defined as symptom resolution. The primary analysis will be a comparison of mean score between treatment arms at day 0, and day 10 of the study. A t-test from an ANCOVA with stratification covariates will be used to assess the mean difference between groups for statistical significance A similar model will be used to analyse the individual domain scores. The proportion of participants in symptom recovery at Day 10 will be compared between treatment groups using the chi-square test.

Symptoms – Common Cold Questionnaire

The Common Cold Questionnaire is a 9 item questionnaire which assess cold symptoms using a 4 point Likert Scale, with scores ranging from 0 to 3. The arithmetic mean of all items will be calculated. The primary analysis will be a comparison of mean score between treatment arms at day 0, day 7, and day 14 of the study. A change of 0.8 [ref] has been used to define the absence of viral symptoms compared to presence of virus and will be used in the analysis.

Blood oxygen saturations

The proportion of days participant had oxygen saturations of <=94% as a proportion of all days monitored within the first 10 days will be compared between the two groups with a Wilcoxon rank sum test. The proportion of participants with at least one assessment <=94% within the first 10 days will be compared using the chi-square test. Exploratory area under the curves for participant blood oxygen saturations will be compared between both study arms using a t-test. Other exploratory time series analysis will also be performed. Amongst participants who did have an abnormal blood oxygen saturation (<=94%), the difference in the rate of recovery will also be assessed.

Body temperature

The proportion of days participant had a body temperature >= 37.5 C as a proportion of all days monitored within the first 10 days will be compared between the two groups with a Wilcoxon rank sum test. The proportion of participants with at least one assessment >=37.5 C within the first 10 days will be compared using the chi-square test. Exploratory area under the curves for participant body temperature will be compared between both study arms using a t-test. Other exploratory time series analysis will also be performed. Amongst participants who did have a raised body temperature (>37.4^{O}C), the difference in the rate of recovery will also be assessed.

SARS-CoV-2 PCR viral load

Changes from baseline to Day 7 and day 14 in SARS-CoV-2 viral load data in the log base 10 scale will be statistically analysed using a linear mixed-effect model or equivalent. The model will contain log base 10 transformed baseline as a covariate, treatment, day, and duration since symptom onset to randomization as fixed effects. The LS means and treatment differences (budesonide minus usual care at each dose level) will be calculated and presented with their corresponding 95% CIs. In addition, the geometric mean ratio to baseline and corresponding standard error for each treatment, and ratio of geometric mean ratio to baseline vs usual care, and corresponding 95% CIs will be presented. All available data will be used in the analysis. The lowest limit of detection for viral load will be assigned if this is negative on PCR.

Missing Data

The following methods of handling missing data can be used.

Non-Responder Imputation (NRI)

For analysis of categorical efficacy and health outcomes variables, missing data will be imputed using an NRI method. Participants will be considered non-responders for the NRI analysis if they do not meet the categorical efficacy criteria or have missing clinical efficacy data at a time point of interest.

Last Observation Carried Forward (LOCF)

A last observation analysis is performed by carrying forward the last postbaseline assessment for the continuous measures or ordinal scale measures. For participants discontinuing the study, the last non missing post baseline observation before discontinuation will be carried forward to the corresponding primary endpoint for evaluation. After LOCF imputation, data from participants with non-missing baseline and at least 1 postbaseline observation will be included in the analyses. These LOCF analyses help ensure that the maximum number of randomized participants who were assessed postbaseline will be included in the analyses.

Mixed-effects Model Repeated Measures (MMRM)

For continuous variables, the primary analysis will be MMRM with the missing-at-random (MAR) assumption for handling missing data. This analysis considers both missingness of data and the correlation of the repeated measurements. No additional imputation methods will be applied to the MMRM analysis.

FLUPRO prespecified missing data analysis

The FluPRO questionnaire has a prespecified method to handle missing data. Please see appendix B for details.

Supplementary/ Additional Analyses and Outcomes

Translational analysis will be undertaken once the samples have been analysed and the data will be combined with the primary and secondary outcome data for analysis. This is detailed in a separate statistical analysis plan and will take place after the main analysis has been reported.

Specification of Statistical Packages

All analysis will be carried out using appropriate validated statistical software such as STATA, SAS, SPLUS or R. The relevant package and version number will be recorded in the Statistical report.

Appendix A

STOIC Study STOP Decision Tree

Scope of this document

1.1. **The purpose of this document is to outline the decision tree for the STOIC study analysis to STOP the study early (stop recruitment, enrolled patients will complete study) or to carry on with the study (continue recruitment)**

1.2. **The decision to consider to stop the STOIC study early has been sought by the CI due to several factors including**

1.2.1. **The waxing and waning nature of recruitment in one region, with expectation that recruitment to sample size will not be complete by February 2020**

1.2.2. **The successful emergence of vaccines targeting SARS-CoV-2**

1.2.3. **Competing studies (PRINCIPLE)**

1.2.4. **These challenging aspects signify concern with recruitment.**

1.3. **Independent statistical advice (StatMind®) will be sought to analyse available data to persuade or dissuade the CI and the study team regarding stopping the STOIC study early**

1.3.1. **The statistical analysis will include estimation of the point estimate and the confidence interval of the primary outcome effect seen in STOIC.**

1.3.2. **Statistical rigour will include**

1.3.2.1. **Simulation to calculate conditional power based on current estimate to achieve a significant outcome at final analysis assuming an early stop. (Note that this final analysis will use all available data for all patients that will complete the study.)**

1.3.2.2. **Analysis of subgroups e.g., selected patients with confirmed positive SARS-Cov-2**

1.3.2.3. **Estimates of effect sizes for selected secondary endpoints**

1.4. **Primary and the selected secondary endpoints for the confidence in the Stop decision are described and defined (see point2 & 3) prior to any data analysis**

Primary Endpoint

1.5. **Urgent healthcare utilisation (includes GP/ED attendances/Hospitalisation)**

1.5.1. **NHS services actively directing to GP and ED simultaneously for COVID since July 2020**

1.5.2. **STOIC protocol powered to show a 50% reduction in urgent healthcare utilisation from 20% in usual care (UC) compared to intervention with inhaled corticosteroid (BUD); alpha 0.05, beta 0.80, n=199 evaluable per group (478 in total when adjusted for 20% withdrawals)**

Secondary Endpoints

1.6. **Protocol defines secondary endpoints as effect of BUD compared to UC in symptoms, physiology and viral inflammation.**

1.7. **Secondary datapoints collected included FluPRO questionnaire, Common Cold questionnaire, Temperature, Oxygen Saturations, Viral load**

1.8. **For the purpose of the Stop decision, only 4 endpoints will be assessed and defined as follows:**

1.8.1. Clinical recovery defined as 'time to symptom resolution, which occurs when daily monitoring ceases in case of recovery.' Patients not in recovery will be censored at Day 29/last day in study if withdrawn. Patients hospitalised will be censored at Day 29.

1.8.2. FluPro questionnaire defined as the change in symptom score between V1 (day 0) and V2 (day 7)

1.8.3. Oxygen saturations defined as the number of days below 94% within 28 days of treatment

1.8.4. Viral load defined as difference between V1 (day 0) and V2 (day 7)

Stop decision tree algorithm

1.9. **Assignation of a positive, neutral and negative effect of BUD compared to UC will be made for the 4 secondary endpoints listed above**

1.10. **A positive effect will be assigned 2; a neutral effect will be assigned 1 and a negative effect will be assigned 0**

1.11. **A stop decision will be made if the total score is 6 or more**

1.11.1. **4 positive effects**

1.11.2. **3 positive effects + 1 neutral**

1.11.3. **2 positive effects + 2 neutral**

1.11.4. **3 positive effects + 1 negative, if negative secondary endpoint is not statistically significant**

Effect size for secondary endpoints; Budesonide vs UC

Variable	Positive	Neutral	Negative
Clinical Recovery	<= -1 day	-1 to +1 days	>= +1 day
FluPro	<-0.3	-0.3 to 0.3	>0.3
Oxygenation	<= -1 day	-1 to +1 days	>= +1 day
Viral Load	<-1 log fold	-1 to +1 log fold	>+1 log fold

Appendix B

Domain	Items	Scoring	Minimum Data Requirement
Nose	Runny or dripping nose Congested or stuffy nose Sneezing Sinus pressure	Arithmetic mean of 4 items within **Nose** domain	Daily score for 3 of 4 items must be present to calculate domain score
Throat	Scratchy or itchy throat Sore or painful throat Difficulty swallowing	Arithmetic mean of 3 items within **Throat** domain	Daily score for 2 of 3 items must be present to calculate domain score
Eyes	Teary or watery eyes Sore or painful eyes Eyes sensitive to light	Arithmetic mean of 3 items within **Eyes** domain	Daily score for 2 of 3 items must be present to calculate domain score
Chest/Respiratory	Trouble breathing Chest congestion Chest tightness Dry or hacking cough Wet or loose cough Coughing Coughed up mucus or phlegm	Arithmetic mean of 7 items within **Chest/Respiratory** domain	Daily score for 5 of 7 items must be present to calculate domain score
Gastrointestinal	Felt nauseous Stomach ache How many times did you vomit? How many times did you have diarrhea?	Arithmetic mean of 4 items within **Gastrointestinal** domain	Daily score for 3 of 4 items must be present to calculate domain score
Body/Systemic	Headache Head congestion Felt dizzy Lack of appetite Sleeping more than usual Body aches or pains Weak or tired Chills or shivering Felt cold Felt hot Sweating	Arithmetic mean of 11 items within **Body/Systemic** domain	Daily score for 8 of 11 items must be present to calculate domain score
Total	All above 32 items	Arithmetic mean of all 32 items within FLU-PRO	In the presence of missing data, the above conditions for the calculation of all domain scores must be met in order to calculate the FLU-PRO total score.

References

1. Powers JH, 3rd, Bacci ED, Leidy NK, et al. Performance of the inFLUenza Patient-Reported Outcome (FLU-PRO) diary in patients with influenza-like illness (ILI). PLoS One 2018;13:e0194180.
2. Powers JH, 3rd, Bacci ED, Guerrero ML, et al. Reliability, Validity, and Responsiveness of InFLUenza Patient-Reported Outcome (FLU-PRO©) Scores in Influenza-Positive Patients. Value Health 2018;21:210-8.
3. Powell H, Smart J, Wood LG, et al. Validity of the common cold questionnaire (CCQ) in asthma exacerbations. PloS one 2008;3:e1802-e.